DOCUMENT RESUME

ED 221 859 CS 207 143

AUTHOR Gebhard, Ann O.
TITLE Quantitative Research in Written Composition.
PUB DATE [74]
NOTE 20p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Computer Oriented Programs; Elementary Secondary Education; *Evaluation Criteria; Evaluation Needs; Higher Education; Literature Reviews; *Writing Evaluation; Writing Instruction; *Writing Research; *Writing Skills; *Written Language

ABSTRACT

Offered as an introductory guide to teachers interested in approaching written English as a "second dialect" that students must master, this review covers quantitative investigations of written language. The first section deals with developmental studies, describing how a variety of researchers have related written structure to writer maturity. Included, among others, are such major works as M. Stormzand and M. O'Shea's 1924 analysis of 10,000 sentences from the writing of students, adults, and professional writers and Walter Loban's longitudinal study, "The Language of Elementary School Children." Research based on the ideas of Noam Chomsky, Kellogg Hunt, and Francis Christensen is also covered. The second section discusses the inherently difficult task of investigating the determiners of quality in composition. Studies by Robert Potter and Margaret Ashida are described in some detail. In the final section, attempts to create computer assessment of student writing are described. (JL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED221859

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Quantitative Research in Written Composition

Ann O. Gebhard, Syracuse University

It has become axiomatic that the child enters school having successfully internalized the rules of his spoken dialect. It is in school that he must master the conventions of what Harold Whitehall (1951) has called "a rather artificial dialect of our language" (p. 101). Graphemes, slowly and often laboriously executed, must replace the phonemes which can be spontaneously produced. The conventions of punctuation and capitalization must come to function in place of the suprasegmentals. Without the help of gesture and voice contour, utterances must by syntactically complete and relatively precise to convey meaning. Roy C. O'Donnell (1967) found definite evidence of syntactical distinctions between the speech and writing of school children appearing as early as grade three.

> Distinct and dramatic differences were found in the syntax of speech and writing in all three grades from which writing samples were collected. On almost all counts, it was clear that where notable differences appeared in Grade 3 (and there were many), they indicated weaker control in writing.
>
> . . . . . .
>
> . . . Unexpectedly uniform evidence, however, showed that advances in the control of syntax in Grades 5 and 7 were accelerated in writing far beyond those reflected in speech (p. 95).

The great task of the child'sschool life is to learn to decode and encode this special dialect; Robert Allen (1966) has likened it to learning "a second language." (p. 349)

The view that written English is a dialect offers a new tool to the teacher of composition. Like the linguistic geographer, he can quantify a description of the grammar of the dialect that interests him. Or, he can measure his students' growth in writing competence against established indices of development. In other words, he and his students can attempt to measure differences in writing style that otherwise might be perceived

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ann O. Gebhard

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS207143

only subjectively.

Thus, research in the quantitative analysis of written composition provides an important tool to the teacher interested in this approach. A number of notable studies have investigated written language in order to define the developmental indices of linguistic maturity. Other research has objectively contrasted writing samples of a lower and higher quality. Still other efforts have attempted to ascertain what measurable writing characteristics approximate human rater judgments to make possible computer evaluation of student essays. The following review is offered as an introductory guide to those teachers interested in the methodology and results of significant, quantitative investigations of written language.

Developmental Studies

In 1917 the influential Hosic report, The Reorganization of English in Secondary School, presaged the future importance of "functional grammar" which was more fully enunciated in the Hatfield Experience Curriculum. Thereafter, studies were designed to discover the types of sentence structure and frequency of error in student writing. The most important of such studies is the earliest. In 1924 M. Stormzand and M. V. O'Shea published the results of their analysis of ten thousand sentences chosen from the writing of students at the elementary, secondary and university level, adult letter and newspaper writers, and professional authors of essays and fiction. No attempt was made to control for mode of discourse. They found that the following structures were important in differentiating more mature from less mature writers:

TABLE I

Stormzand O'Shea Analysis of Developmental Writing Patterns

	(1000) (sentences) Fourth Grade	(1000) (sentences) High School	(1500) (sentences) University Average	(6500) (sentences) Adult Average	(500) (sentences) Robert Louis Stevenson
Mean Sentence Length	11.1	18.2	20.7	20.9	31
Number of Dependent Clauses per 100 Sentences	24	79	119	97	143
Ratio of Prepositional Phrases per Sentence	.85	2.15	2.15	2.45	3.89
Ratio of use of Verbals per Sentence	.23	.68	.74	.77	.89

The type of clause used, adjectival, adverbial or noun, was distributed about equally throughout the material, but the use of compound structures declined from a high of 25% in the fourth grade to an adult average of 17.2%. A steady incremental growth in sentence complexity, as evidenced by increased use of phrasal and clausal structures, had been shown to parallel increased maturity.

But the Danish linguist, Otto Jesperson (1927), theorized that women use subordination less often than men. In discussing "the linguistic symptoms of a peculiarity of feminine psychology which has not escaped observation," Jesperson says,

> If we compare long periods as constructed by men and by women we shall in the former find many more instances of intricate or involute structures with clause within clause, a relative clause in the middle of a conditional clause or vice versa with subordination and sub-subordination, while the typical form of long feminine periods is that of coordination, one sentence or clause being added to another on the same plane and the gradation between respective ideas being marked not grammatically, but emotionally, by stress and intonation and in writing by underlining. (p. 251)

Such a contention was disproved by Lou La Brant. Sex differences in her sample of 924 compositions of Kansas elementary and high school students were insignificant except in fluency, boys' papers averaging 136 words while girls' averaged 163 words. La Brant formulated an index of written language development used by later researchers, the "subordination ratio," obtained by dividing the number of subordinate predicates by the total number of predicates used. Incremental gains in subordination ratio held steady over the school years following chronological age more closely than mental age; the .11 subordination ratio of eight-year old developed into the .36 of the sixteen-year old. Although the average dependent clause length remains

fairly constant (7.2 words in the elementary school and 8.0 words in the high school) substantial increases in the subordination ratio (.46) and average clause length (15.2 words) characterized the published writing of twenty-one psychologists which La Brant also analyzed.

In her study of the compositions of 959 Minnesota students written the same year, Ellen Frogner, like Stormzand and O'Shea and La Brant, found a pattern of growth in the increased use of complex structures by older students. She also found that participial phrase use increased proportionately to grade level: 3.1% of the sentences in grade seven contain a correctly used participle; 4.4% in grade nine; and 6.5% in grade eleven. La Brant had commented on increased use of infinitives but had not found it to be a constant corollary of development. La Brant had noted that dependent clauses not only increased in frequency with greater maturity but also "increased in complexity and clarity of thought."

During the late forties and fifties research interest in quantitative measures of composition maturity declined. A recent essayist (Myers, 1971) has remarked that the metaphoric model for the English program during these decades was, respectively, "English as Personality Adjustment," and "English as Social Studies." This perhaps accounts for the dearth of objective studies.

The publication in 1963 of the first NCTE research monograph, Walter Loban's longitudinal study, The Language of Elementary School Children, seems to signal a revival of interest. Although primarily concerned with oral language, Loban helped to establish a methodology of significance to future composition research. He found "that not pattern but what is done to achieve flexibility within the pattern proves to be a measure of effective-

ness and control of language at this level of development." Loban endorsed transformational grammar, which he had used in a very limited way, as "a more precise method of measuring grammatical complexity." (p. 63)

Structural linguistics had provided an elaborate phonological system for the classification of language sounds, but was less thorough in its approach to the analysis of written language. With the publication of Syntactic Structures (1957), Noam Chomsky redirected the interest of linguists to their traditional concern with syntax. He addressed himself not so much to the problem of classification of existing structures, but rather to the question of the generation of structure. One might say that attention transferred from the process of "decoding" to the process of "encoding." His theory that sentences are generated from "deep structure" thought processes to "surface structure" syntactical and phonological or graphemic forms by the dynamics of "transformations" has been crucial to the development of more exact methods of written language research.

Extremely influential has been Kellogg Hunt's pioneer work of transformational analysis of student writing, Grammatical Structures Written at Three Grade Levels (1965), in which he established as an index of stylistic maturity the average length of T-unit, the shortest grammatically allowable unit into which sentences can be divided, a measure similar to Loban's "communication unit." Hunt based his contrastive analysis chiefly on the age of the student writers, analyzing 100-word samples written by fourth, eighth, and twelfth graders of average ability.

In his study Hunt found that as writers matured they use fewer simple sentences of identification. (My dog's name is Jimmy. He is brown and he has spots) and more structures that combine and consolidate meaning (I own

a spotted, brown dog named Jimmy). It is this factor that accounts for the consistent lengthening of T-unit, the base sentence with its clause modifiers. Fourth grader's T-units average 8.6 words; eighth graders', 11.5 words; and twelfth graders', 14.4 words. Hunt rejects sentence length, which had elsewhere been cited as an index of maturity, because of the tendency of young writers to string together a succession of simple sentences with the coordinating conjunction and. Thus the key question in Hunt's study becomes how mature writers secure the greater expansion of their T-units. His answer:

> The chief factor which lengthens clauses appears to be the increasing of nonclause modifiers of nouns and the nominalization of clauses. This factor and the increase in adjective clauses account in the main for increased length of T-unit. . . .
>
> The older student can incorporate and consolidate more grammatical structures into a single grammatically related unit. The younger student produced short separate units. As he matures that span broadens, so he casts the net of consolidation over larger and larger bodies of material. As he consolidates, he also discards needless words. His redundancy lessens and his succinctness gains. (p. 143)

A more recent study which has emphasized such sentence combining transformations is an analysis of children's development of language mastery is Syntax of Kindergarten and Elementary School Children: A Transformational Analysis (1967) written by Roy C. O'Donnell, William J. Griffin, and Raymond Norris. This work concerns itself with oral as well as written language and corroborates Hunt's conclusion that T-unit length is a valid measure of increasing linguistic maturity.

As the subjects' grade level advanced, the researchers found regular, statistically significant increases in the use of sentence combining transformations involving coordinate constructions within clauses, subclausal adverbial constructions, and of nominal constructions containing adjectives, participles, and prepositional phrases. They speculate that increasing use

of deletion transformations may be an important indication of increasing maturity.

Interesting in terms of the theory that written English is a special dialect of spoken English is the finding of these investigators that the syntax of the writing of the fifth and seventh graders of the study, the two oldest groups analyzed, differed materially from their speech in that the writing samples contained a fuller range of grammatical structures than those utilized in oral expression.

Even though O'Donnell's study supported Hunt's theory of the validity of T-unit length as a measure of writing maturity, rhetorician Francis Christensen (1968) has attacked the notion of such an index, claiming that it may reflect increasing age and grade but does not assess skill in writing. Christensen has suggested that Hunt's continuum is not necessarily one of improvement -- and that his twelfth graders, even if they continue to grow in terms of length of T-unit are never going to write like skilled adults. Undoubtedly the most influential figure of his time interested in the quantitative analysis of prose, Christensen offered provocative arguments to counter Hunt's claim. His thesis, put forward in Notes Toward a New Rhetoric (1967) is that the most important element of the style of mature fiction writers like Hemingway or Faulkner, as well as non-fiction writers, is their use of what he called the "cumulative sentence," a sentence with a short base clause in which most of the author's meaning is conveyed by the use of modification that is loose and additive, and in the final sentence position such as absolute phrases, non-restrictive clauses, and appositive, verbal, adjective and adverb phrases. Christensen feared that studies such as Hunt's and O'Donnell's would persuade school men to teach their students

to write with increasing complexity, thus encouraging the writing of stilted, involuted prose -- what he calls "pretzel prose." In fact it was a doctoral study (Mellon, 1970) testing the validity of sentence-combining practice as a way of increasing students' syntactic control which prompted Christensen's criticism of Hunt's index. Although well-meaning, Christensen's remarks seem to be based partially on a misreading of Hunt, who like Christensen, finds increasing use of structures of modification crucial to increased T-unit length. The rhetorician may be very correct in his assumption that the average ability twelfth graders of Hunt's study will never write like skilled adults: some structures which Christensen's research has found popular in professional writing, the absolute phrase, for instance, are not used at all in the reported samples of student writing; however, this assumption does not lessen the importance of studies such as Hunt's which try to trace the development of stylistic competence.

One recent researcher (Wolk, 1970) tested Christensen's premise that final free modification is the mark of a professional writer by analyzing a group of twelve essays in magazines, such as Harper's and the Saturday Review, and comparing the results with college student writing. He found that 20.9% of student writing occurs in structures of free modification as compared to 31.0% of the professional. While the magazine writers use free modification more frequently in the initial and medial (after the complete subject but before the end of the complete predicate) position than the students, reflecting their greater overall use, the greatest degree of contrast in the use of such structures occurs in the final position, 8.4% for the students to 13.7% for the professional -- a finding that supports Christensen's assumption.

An earlier investigator was particularly interested in the way student writers and professionals begin their sentences. Herman Struck (1965) contrasted the writing of graduate students with that of published writers. He found that over fifty per cent of the published sentences begin with the subject; non-subject openers were used as follows: dependent clauses, 6%; prepositional phrases, 13%; adverbs, 8%; coordinate conjunctions, 12%; and verbals and expletives, 6%. Although these figures differ considerably from Christensen's, the findings are provocative, particularly in terms of the high incidence of professional sentences that begin with coordinating conjunctions, a form frequently condemned in secondary English classrooms. Christensen's assessment of the sentence beginnings of a group of distinguished contemporary writers found they began sentences with constructions other than the subject only 25%of the time, but his method of counting may have differed materially from Struck's.

As Table 2 indicates, a number of reliable indices of language development in terms of chronological maturation have been established. But when one considers style in writing -- subjectively perceived "good" writing as opposed to "poor" writing -- how helpful are these indices? Or, in other words, to what extent do the indices of writing maturity function as indices of writing quality?

Investigation of the Determiners of Quality in Composition

An English investigator attempted to obtain a factor analysis of the determiners of quality-rated composition. After judges had rated the themes of "some hundreds" of children on the topic, "How the Twins Found the Treasure," Percival (1963) analyzed them in terms of number of words, average length of sentence, number of coordinating conjunctions, different kinds of punctuation,

TABLE 2

Developmental Indices of Writing Maturity

Age or Grade Level	3/8-9	4/9-10	5/10-11	6/11-12	7/12-13	8/13-14	9/14-15	10/15-16	11/16-17	12/17-18	College Students	Superior Adults
Mean sentence length											20.7	
Stormzand & O'Shea			11.1									31
Davis (mean of 5 longest sentences)		15.6										
Heider & Heider Hearing children	10.2	11.1	11.1	12.8	13.7	13.9						
Deaf "			7.9	8.0	9.0	8.4	9.6	9.5	10.2			
Mean clause length												
La Brant		7.2							8.0			15.2
Hunt		6.6				8.1				8.6		11.5
O'Donnell	6.5		7.4		7.7							
Subordination ratio												
La Brant		.11							.36			.46
Anderson											.50	
Mean T-unit length												
Hunt		8.51				11.34				14.4		20.3
O'Donnell	7.67		9.34		9.99							

subordinating conjunction, effective language, accuracy of composition, and accuracy of spelling. He concluded that "the most predictive determiner (of quality) was the number of w ds in a composition!" (p. 120)

There are obviously formidable barriers to research endeavoring to identify valid quantifiable measures of prose quality. The difficulties of trying to translate subjective evaluation into objective criteria are obvious. Among the variables that effect measurement of writing ability discussed by Braddock, Lloyd-Jones and Shorer (1963) in their authoritative Research in Written Composition are writer, assignment, rater and colleague.

The degree to which mode of discourse, an element of the assignment variable, influences syntactic complexity has been demonstrated by Christine San Jose (1972) in her study Structures in Four Modes of Writing at Fourth Grade Level.

While this consideration is relevant to any quantitative assessment of writing, it seems more crucial in its effect upon judgment of writing quality. Suppose the gifted writers in Percival's sample chose to develop composition topic in a narrative mode, which in terms of San Jose's research evidences more simple structures, while the poor writers choose exposition, which generally manifests more complex structures -- what effect would this have on quantitative results?

In 1966, Robert Potter attempted to quantify the differences in contrasting samples of quality-rated tenth grade compositions. He secured five-hundred-word compositions on the topic, "The Qualitites of a Good Teacher," from one hundred students, and after correcting gross mechanical errors, had them re-typed. Four experienced English teachers were asked

to mark each paper "good," "average," or "poor." Subsequently, two additional raters confirmed the initial rankings, and thus the twenty best and the twenty worst papers of the group were identified. The examination and comparison of these papers form the basis of Potter's study. Briefly, these were the results of his contrastive analysis of the "good" and "poor" writing samples.

1. The good papers displayed more variety in coordination and subordination devices and employed more transitional devices.
2. The good papers employed more words in structures of modification, particularly prepositional phrases.
3. Poor papers evidenced more nominal clause construction than did good papers and these clauses were more likely to omit the introductory "that." (This was largely explained by the poor writer's tendency to begin his sentences with terms like "I guess," "I think," etc.)
4. Sentences on the good papers began with twice as many verbal structures in the pre-subject position as did sentences on the poor papers.
5. Lou La Brant's subordination ratio was exactly the same in the two groups.
6. Both sentence length and T-unit length were greater for the good papers (17.8 - 16.0) than for the poor papers (15.9 - 14.2 respectively). Since Potter did not test for significance, it is difficult to determine how meaningfulthese figures are.

Writing in 1967, Margaret Ashida produced a thorough-going study in which she contrasted the writing of freshman composition students and

professionals. (Her method of choosing the professional sample is outlined elsewhere in this paper.) Her freshman sample was based on the writing of twenty-five randomly selected University of Michigan students. The purpose of her research seems to be an exercise in practical pedogogy; she wanted "to identify the areas of language activity which have the most functional value." (p. 1) Earlier she and Leslie T. Whipp (1963) had reported favorably on a composition course they taught at the University of Nebraska in which students were required to make extensive quantitative analyses of their own prose as well as published essays.

Her research challenges the assumptions of both Christensen and Hunt. She found that over 35.14 of the professionals' sentences began with openers as contrasted with 30.4% of freshman sentences. She ascribes Christensen's lower figure to the biased nature of his sample. (He chose professional writers he liked.) Hunt's T-unit also comes in for censure in this study, as inadequate in reflecting the stylistic subtleties of professionals who often use sentence fragments deliberately. She would substitute the "Ellie," "a word or group of words which either serves as or comprises an independent clause plus all grammatically subordinate elements, if any." (p. 166)

Ashida reports significant differences not only in terms of length of utterance between freshman and professionals: sentence length, 24.49 words for professionals to 19.45 for the freshman, and clause length, 12.1 words to 10.5; but also in terms of sentence structures: professionals use more complex and compound-complex sentences than do the freshmen, having a subordination ratio of 41.4 to the freshmen's 38.2. Ashida concluded, "Investigations of form -- the length, range, and standard deviation of all sentences . . . consistently showed that F-group (freshman) structures are

deficient in both length and variety." (p. 307)

Addressing themselves to the question, "In what ways may measurable differences that generally distinguish the language performance of older children from that of younger children be related to distinctions between better and poor writing elicited under uniform stimulus conditions from children in a particular age range?", Sister Jeanne Marie Jurgens and William J. Griffin (1970) analyzed 299 quality rated compositions written by Tennessee secondary students in terms of seven linguistic features. Again, compared to the developmental studies, their results were indefinite. Although the quality sub-groups were distinguished by increases in total words, total T-units, and total clauses, neither T-unit length nor clause length, which have elsewhere been established as reliable developmental measures, differentiated among the groups.

When Ramon Veal (1974) analyzed a smaller sample of the quality-rated writing of 81 Georgia school children in grades two, four, and six, he found that T-unit length proved to be a most effective syntactic marker of quality." Total words in a composition and the total number of subordinate clauses rather than mean clause length seemed also to be reliable correlates of high quality in this study.

Computer Evaluation of Composition Quality

While it seems obvious that the degree of uniformity present in the developmental studies does not exist in the qualitative studies, a more definite, if somewhat arbitrary, approach to the problems of stylo-statistics has been offered by computer science. Slotnick (1972) asserts unequivocally, "There is a statistically identifiable relationship between essay grades assigned by human judges and the measures computers can make." (p. 253)

The pioneer work in this field was done by Page and Paulus (1968) who endeavored to create a computer program which would simulate the behavior of human judges. The researchers theorized that for a given intrinsic quality, such as diction, essays would reflect certain measurable characteristics, such as the proportion of words not on a common word list, high mean word length or standard deviation of word length. The computer assessment of these quantifiable variables would approximate the human graders' subjective evaluation of the comparable quality. Using a multiple-regression analysis these objective "proxes" were correlated with the ratings of experienced judges of one thousand student compositions written at the University of Wisconsin campus school. The judges were asked to rate the essays on content, organization, style, mechanics, and creativity. The most important "prox" turned out to be "length of essay" which correlated with every one of the evaluative areas except, as might be logically expected, mechanics. Standard deviation of word length figured importantly in every area. Interestingly enough, more arbitrary "proxes" such as frequency of commas and dashes and the presence of a title also had high correlation with graders' judgments. In all, thirty proxes were established which produced "quite humanoid ratings of overall quality," (p. 165) with a predictive value of .53 compared to the human judges predictive rating of .54. A human rater reliability of .5 is typical in subjective assessment.

Whelan (1971) posited another model after he entered a sample of seventy-one previously graded seventh grade essays into the computer. For these essays the following quantifiable variables, taken together, had the predictive worth of .6: number of capitalization errors (negative

value), standard deviation of sentence length, standard deviation of word length, number of connectives, occurrence of "then," average sentence length, and number of usage erros (negative value). This model illustrates two characteristics of computer assessment of composition. It tends to be specific to the compositions on which it is programmed; this topic involved chronology, hence the positive effect of the adverb "then." It is also limited by the ingenuity of the programmer. In order to correct capitalization errors, the computer must be provided with a list of the proper nouns that would probably be used in the student essays. Creating such a list for more mature writers, composing themes on a wide range of topics, might be prohibitively difficult.

Yet another program was developed by Slotnick (1972) using factor analysis to isolate the quantifiable variables that would approximate the more intrinsic values of fluency, diction, sentence structure, and paragraphing. His model was also interested in punctuation and spelling. Again, total words in a composition was highly predictive of quality. Mean sentence length and standard deviation of sentence length, as might be expected, were closely related to sentence structure. This researcher (1971) has emphasized that computer evaluation, even with more sophisticated techniques than presently exist, should be considered as an adjunct rather than a substitute for the sensitivity and perceptiveness of the human judge.

Traditionally, composition quality has been evaluated only subjectively so that even the earnest student has difficulty ascertaining clear and non-controversial ways of improving his writing. Research studies such as those described in this review may help point the way to more specific, realizable goals for writing programs.

REFERENCES

Allen, Robert L. "Written English is a second language." In Gary Tate and Edward P. J. Corbett, (Eds.) Teaching high School composition. New York: Oxford University Press, 1970.

Anderson, John E. "An evaluation of various indices of linguistic development." Child Development, 1937, 8, 62-68.

Ashida, Margaret E. "From, syntax and statistics: a quantitative approach to written composition." Unpublished Ph.D. dissertation, University of Nebraska, 1967.

Ashida, Margaret E. and Leslie T. Whipp. "A slide rule composition course." College English, September 1963, 26, 18-22.

Braddock, Richard, Richard Lloyd-Jones and Lowell Schoer. Research in written composition. Champaign, Illinois: National Council of Teachers of English, 1963.

Christensen, Francis. Notes toward a new rhetoric. New York: Harper and Row, 1967.

Christensen, Francis. "The problem of defining a mature style." English Journal, April 1968, 57, 574-578.

Chomsky, Noam. Syntactic structures. The Hague: Mouton, 1957.

Davis, Edith A. "Mean sentence length compared with long and short sentences as a reliable measure of language development." Child Development, 1937, 8, 69-79.

Frogner, Ellen. "Problems of sentence structure in pupils' themes." English Journal, November 1933, 22, 742-49.

Heider, F. K. and G. M. Heider. "A comparison of sentence structure of deaf and hearing children." Psychological Monographs, 1940, 52, 42-103.

Hunt, Kellogg W. Grammatical structures written at three grade levels. Champaign, Illinois: National Council of Teachers of English, 1965.

Jespersen, O. Language: its nature, development and origin. London: Allen & Unwin, 1922.

Jurgens, Sister Jeanne Marie and William F. Griffin. The relationship between overall quality and seven language features in composition written in grades seven, nine and eleven. Nashville, Tenn." George Peabody College for Teachers, 1970.

LaBrant, Lou L. "A study of certain language developments of children in grades 4-12 inclusive." Genetic Psychology Monographs, November 1933, 14, 387-491.

Loban, Walter D. The language of elementary school children. Champaign, Illinois: National Council of Teachers of English, 1963.

Mellon, John C. Transformational sentence combining: a method for enhancing the development of syntactic maturity. Champaign, Illinois: National Council of Teachers of English, 1968.

O'Donnell, Roy C., William J. Griffin, Raymond C. Norris. Syntax of kindergarten and elementary school children: a transformational analysis. Champaign, Illinois: National Council of Teachers of English, 1967.

Page, Ellis and Dieter Paulus. The analysis of essays by computers. Storrs, Connecticut: The University of Connecticut, 1968.

Percival, E. "A quality scale of English composition for an 11 plus age group." Unpublished M.S. thesis, University of Birmingham, 1958, cited by Edwin A. Peel in "The analysis of preferences." Research Design and the Teaching of English. Champaign, Illinois: National Council of Teachers of English, 1963.

Potter, Robert. "An exploratory study of the relationship between certain aspects of sentence structure and the overall quality of tenth grade writing." Unpublished Ph.D. dissertation, Columbia University, 1966.

San Jose, Christine Martinez. "Grammatical structures in four modes of writing at fourth grade level." Unpublished Ph.D. dissertation, Syracuse University, 1972.

Slotnick, H. B. "Toward a theory of computer essay grading." Journal of Educational Measurement, 1972, 9, 253-63.

Slotnick, H. B. and J. V. Knapp. "Essay grading by computer: a laboratory phenomenon." English Journal, February 1971, 60, 75-80.

Stormzand, M. and M. J. O'Shea. How much English grammar? Baltimore: Warwick and York, 1924.

Struck, Herman R. Better prose: a method. Boston: Houghton, Mifflin Company, 1965.

Veal, L. Ramon. "Syntactic measures and rated quality in the writing of young children." Studies in Language Education. Athens: University of Georgia, 1974.

Whitehall, Harold. Structural essentials of English. New York: Harcourt, Brace and World, Inc., 1951.

Wolk, Anthony. "The relative importance of the final free modifier: a quantitative analysis." Research in the Teaching of English, 1970, 4, 59-68.